# STATISTICAL PLANT SET ESTIMATION USING SCHROEDER-PHASED MULTISINUSOIDAL INPUT DESIGN

D. S. Bayard
Jet Propulsion Laboratory
California Institute of Technology
4800 Oak Grove Drive
Pasadena, CA 91109

## ABSTRACT

In this paper, a frequency domain method is developed for plant set estimation. The estimation of a plant "set" rather than a point estimate is required to support many methods of modern robust control design. The approach here is based on using a Schroeder-phased multisinusoid input design which has the special property of placing input energy only at the discrete frequency points used in the computation. A detailed analysis of the statistical properties of the frequency domain estimator is given leading to exact expressions for the probability distribution of the estimation error, and many important properties. It is shown that for any nominal parametric plant estimate, one can use these results to construct an overbound on the additive uncertainty to any prescribed statistical confidence. The "soft" bound thus obtained can be used to replace "hard" bounds presently used in many robust control analysis and synthesis methods.

## 1. INTRODUCTION

The goal of robust control design is to synthesize a controller which establishes certain closed-loop properties (e.g., stability, performance, sensitivity reduction, etc. ) for a specified set of open-loop plants. The set of open-loop plants is typically characterized using a-priori information concerning the physics of the 'system, system modelling, engineering judgement, experience with similar systems, etc..

In the interest of reducing conservatism in the plant uncertainty description, there has been recent efforts aimed at characterizing the plant set using system identification techniques [6][14][15][17][18][22][25][26]. In the case that experimental input/output data is available from the system, this requires characterizing the set of plants which are consistent with (or equivalently, can't be discounted based on), the data. In order to best support robust control objectives, it is of interest to find a precise statistical characterization of the frequency domain uncertainty. For this purpose, the present paper develops a frequency domain estimation method based on a Schroeder-phased sinusoidal input design, The usefullness of the Schroeder-phased input design for plant set estimation follows from certain key properties of the error distributions established in this paper,

- The error distributions are plant independent

- The complete error probability distributions are available rather than just the means and covariances

- The error probability distributions are *exact* and are not asymptotic approximations

1

- The DFT estimator is unbiased for any data length and hence does not require special frequency windowing functions

- The errors at each frequency grid point are statistically independent

It is shown that these properties lead to a precise characterization of the plant set toa specified statistical confidence, e.g., $(1 - \alpha) \cdot 100\%$. The significance of this result is that if a robust controller is designed to provide some specified level of stability or performance *for all plants lying in the* additive *uncertainty set*, then with probability $1 - \alpha$ the controller will work as planned *when applied to the true system.*

Aside from robust control applications, one of the most important practical properties listed above, is that the DFT estimator gives an unbiased estimate, i.e., *the frequency domain estimate* is *free from* windowing distortions. This may seem somewhat remarkable to researchers familiar with the usual "leakage effects" associated with using white noise, burst-random, or pseudo-binary random inputs. Details of the implementation required for distortionless frequency domain estimation are contained in the paper.

There is much to be gained from using statistical uncertainty characterizations since they arc potentially less conservative than deterministic uncertain y characterizations. For example, the notion that noise disturbances tend to "average out" over time is completely missing from deterministic treatments (cf., Helmicki, Jacobson and Nett [15]). *On* the other hand, there is very little to lose from using a statistical approach i.e., a controller based on hard bounds can be designed to work with probability 1, while a controller based on statistical bounds can be designed to work with, for example, probability .9999.

## 2. STATISTICAL ADDITIVE UNCERTAINTY BOUNDS

### 2.1 Background

For this discussion, it will be useful to make the following assumptions,

**Assumption 1** The true plant is a single-input, single-output unknown exponentially stable linear time-invariant (LTI) transfer function assumed to have a sampled-data representation $P^*$ *(Z-l)* in the shift operator z-l. ∎

**Assumption 2** The output disturbance v(k) can be represented by v(k)= $Wd(k)$ where $d(k)$ is a white Gaussian zero-mean noise sequence normalized such that $E[d(j)d(k)] = \delta_{jk}$; $W$ is a linear filter which be decomposed as $W(z^{-1}) = \sigma\overline{W}(z^{-1})$ where $\sigma <$ 00' is a scalar (possibly unknown); and $\overline{W}(z^{-1})$ is a *known* stable and stably invertible transfer function. •

An additive error $\Delta_A(z^{-1})$ is used to characterize the mismatch between the true plant $P^*(z')$ and a nominal plant estimate $P^o(z^{-1})$, i.e.,

$$\Delta_A(z^{-1}) = P^*(z^{-1}) - P^o(z^{-1}) \tag{2.1}$$

It will be useful to define the set of plants $\Omega_A(P^o, \ell_A(\omega))$ associated with a specified overbound $\ell_A(\omega)$ *on* the additive error, i.e.,

$$\Omega_A(P^o, \ell_A(\omega)) = \{P : |P - P^o| \le \ell_A(\omega), \text{ for all } \omega \in [0, \pi/T]\} \tag{2,2}$$

This notion is extended to the specification of a *statistical* overbound $\ell_A^{1-\alpha}(\omega)$ in the following definition,

**Definition 2.1** $\ell_A^{1-\alpha}(\omega)$ is said to bean *overbound on the additive uncertainty with statistical confidence* $(1 - \alpha)$ x 100% if,

$$Prob\{P^* \in \Omega_A(P^o, \ell_A^{1-\alpha}(\omega))\} \geq 1 - \alpha \qquad (2.3)$$

∎

The significance this definition is that $\ell_A^{1-\alpha}$ characterizes (to statistical confidence $(1 - \alpha)$ x 100%) a set in which the true plant belongs. Hence, if a robust controller is designed to provide some specified level of performance for *all plants lying* in *the additive uncertainty set* $\Omega_A(P^o, \ell_A^{1-\alpha}(\omega))$, then with probability $1 - \alpha$ the controller will work as planned *when* applied *to the true system.*

A main result of this paper will be a systematic method for determining $\ell_A^{1-\alpha}$ directly from experimental data. In anticipation of this result, let $\chi$ denote a set of experimental data taken on the system. The specific form of this data will be discussed in more detail later. Assume that a nominal stable parametric plant estimate $P^o$ is fit to the data. Given $\chi$ and $P^o$, we define the following mapping.

**Definition 2.2** $B^{1-o}$-rule:

$$B^{1-\alpha} : \chi, P^o \rightarrow \ell_A^{1-\alpha}(\omega) \qquad (2.4)$$

where $\ell_A^{1-\alpha}(\omega)$ is a nonparametric overbound on the additive uncertainty with statistical confidence $(1 - \alpha)100\%$.

m

Intuitively, for any nominal plant estimate $P^o$, the $B^{1\cdot\alpha}$-rule is a rule for determining the statistical additive uncertain y overbound $\ell_A^{1-\alpha}(\omega)$ directly from experimental data $\chi$. The interpretation is that of characterizing (to a specified statistical confidence) the "ball" of plants about the nominal estimate $P^o(z^{-1})$ *which cannot be discounted based on the data.*

## 2.2 Constructing $B^{1-\alpha}$-Rules

A general method to construct $B^{1\,\text{''}}$-rules for sampled-data systems from noisy frequency domain data will be discussed in this section. The bounds derived here can then be transformed into the continuous domain (or more precisely the w-plane) using a Tustin transformation for use with modern robust control design software [2][9].

**Definition 2.3** A discrete-time MIMO transfer function $G(z^{-1})$ is said to be in $\mathcal{D}(M, \rho)$ if the impulse response sequence $\{g(kT)\}_{k=0}^{\infty}$ defined by $g = \mathcal{Z}^{-1}(G(z^{-1}))$ satisfies $|g(kT)| \leq M\rho^k$ for some $\infty > M > 0$ and $1 > \rho \geq 0$. •

Lemma 2.1

Let $G(z^{-1}) \in \mathcal{D}(M, \rho)$. Then the modulus of the derivative of G on the unit circle can be uniformly bounded from above as follows,

$$\left\| \frac{dG(e^{-j\omega T})}{d\omega} \right\|_I < \frac{TM\rho}{(1-p)}, \qquad (2.5)$$

**Proof:** The result follows from differentiating $G(e^{-j\omega T}) = \sum_{k=0}^{\infty} g(kT)e^{-j\omega k T}$ and using the fact from Definition 2.3 that $|g(kT)| \leq M\rho^k$. Details can be found, for example, in [6]. ∎

A general result for constructing $\mathcal{B}^{1-\alpha}$-rules from frequency domain data is given next.

**Theorem 2.1**

Given discrete-time plant $P^*(z^{-1}) \in D(M, p)$, assume that noisy frequency domain data $\{\hat{P}(\omega_i)\}_{i=1}^{N}$ are available on a uniform grid on the unit circle $\omega_i = i\Delta_g, i = 1, ..., N$ with grid spacing $\Delta_g = \omega_{i+1} - \omega_i = \frac{\pi}{TN}$. Assume that the accuracy of each data point can be characterized by the quantity $\epsilon_i$ such that the event $E_i$,

$$E_i . \quad |P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)| \leq \epsilon_i \tag{2.6}$$

is satisfied with at least probability $1 - \kappa$ at each grid point $i$. Here, the events $E_i\, i = 1, ..., N$ may or may not be jointly statistically independent. Let $S(\hat{P}, \omega)$ be a linear spline interpolant to the data $\{\hat{P}(\omega_i)\}_{i=1}^{N}$, i.e.,

$$S(\hat{P}, \omega) = \begin{cases} \hat{P}(\omega_i) + \frac{(\omega - \omega_i)}{\Delta_g}(\hat{P}(\omega_{i+1}) - \hat{P}(\omega_i)) & \text{for } \omega \in (\omega_i, \omega_{i+1}] \\ \hat{P}(\omega_1) & \text{for } \omega \in [0, \omega_1]. \end{cases} \tag{2.7}$$

and let $P^o(z^{-1})$ be any stable parametric model fit to the data. If $\ell_A^{1-\alpha}(\omega)$ is defined as,

$$\ell_A^{1-\alpha}(\omega) = B_1(\omega) + B_2(\omega) \tag{2.8}$$

where,

$$B_1(\omega) = \begin{cases} \epsilon_{i+1} + |S(\hat{P}, \omega) - \hat{P}(\omega_{i+1})| + \frac{\Delta_g T M \rho}{(1-\rho)^2} & \text{for } \omega \in (\omega_i, \omega_{i+1}] \\ \epsilon_1 + \frac{\Delta_g T M \rho}{(1-\rho)^2} & \text{for } \omega \in [0, \omega_1] \end{cases} \tag{2.9}$$

$$B_2(\omega) = |S(\hat{P}, \omega) - P^o(e^{-j\omega T})| \tag{2.10}$$

then, $\ell_A^{1-\alpha}$ is an *overbound* on the additive *uncertainty with statistical confidence,*

$$1 - \alpha = \begin{cases} (1 - \kappa)^N & \text{if } E_i, i = 1, ..., N \text{ are independent} \\ 1 - \kappa N & \text{otherwise} \end{cases}$$

*i.e.,* $\ell_A^{1-\alpha}(\omega)$ *satisfies,*

$$Prob\{P^* \in \Omega_A(P^o, \ell_A^{1-\alpha}(\omega))\} \geq 1 - \alpha \tag{2.11}$$

**Proof:** Consider the following inequality,

$$|P^*(e^{-j\omega T}) - P^o(e^{-j\omega T})| \leq |P^*(e^{-j\omega T}) - S(\hat{P}, \omega)| + B_2(\omega) \tag{2.12}$$

where $B_2(\omega) = |S(\hat{P}, \omega) - P^o(e^{-j\omega T})|$. The term $B_2$ can be computed directly for all $\omega$. The 1st term on the right hand side of (2.12) can be bounded as follows for $\omega \in [\omega_i, \omega_{i+1}]$,

$$|P^*(e^{-j\omega T}) - S(\hat{P}, \omega)| = \left| P^*(e^{-j\omega_{i+1}T}) - S(\hat{P}, \omega_{i+1}) - \left( S(\hat{P}, \omega) - S(\hat{P}, \omega_{i+1}) \right) \int_\omega^{+1} \frac{dP^*}{d\omega} d\omega \right|$$

$$\leq \epsilon_{i+1} + |S(\hat{P}, \omega) - S(\hat{P}, \omega_{i+1})| + \int_\omega^{\omega_{i+1}} |\frac{dP^*}{d\omega}| d\omega$$

$$\leq \epsilon_{i+1} + |S(\hat{P}, \omega) - \hat{P}(\omega_{i+1})| + \frac{\Delta_g T M \rho}{(1-\rho)^2} \triangleq B_1(\omega) \qquad (2.13)$$

Here, the last inequality follows from (2.7) and the bound given by Lemma 2.1. on the transfer function derivative $dP^*/d\omega$. From the above analysis it follows that $\ell_A^{1-\alpha}(\omega)$ is an overbound on $|P^*(e^{-j\omega T}) - P^o(e^{-j\omega T})|$ for all $\omega \in [0, \pi/T]$ if events (2.6) are satisfied at all grid points simultaneously.

If events $E_i$, $i = 1,..., N$ are statistically independent, and each has probability $1 - \kappa$ of being true, the probability of all events being satisfied simultaneously is given by $1 - \alpha = (1 - \kappa)^N$. Alternatively, 'if the events $E_i, i = 1,..., N$ are not statistically independent, let $\overline{E}_i$ denote the complementary event of $E_i$. Then one can derive a useful Bonferroni inequality (cf., Feller [10], pp. 110) as follows,

$$Prob\left[ \bigcap_{i=1}^{N} E_i \right] = 1 - Prob\left[ \bigcup_{i=1}^{N} \overline{E}_i \right] \geq 1 - \sum_{i=1}^{N} Prob[\overline{E}_i] = 1 - \kappa N \qquad (2,14)$$

In this case, the probability of all events (2.6) being satisfied simultaneously is seen from (2.14) to be at least $1 - \alpha = 1 - \kappa N$. ∎

As desired, $\ell_A^{1-\alpha}(\omega)$ in Theorem 2.1 defines a $\mathcal{B}^{1-\alpha}$-rule since it is a function only of the nominal plant $P$" and the experimental plant data set $X = \{M, p, \{\hat{P}(\omega_i), \epsilon_i\}_{i=1}^{N}\}$. Intuitively, the overbound $\ell_A^{1-\alpha}(\omega)$ in (2.8) can be thought of as the sum of three terms: a *curve fit* error $B_2(\omega)$; an *estimation error* at the grid points $\epsilon_i$; and an *interpolation error* between grid points $B_1(\omega) - \epsilon_i$.

Values for $M$ and $p$ will be assumed known a-priori (they maybe known from the physics of the process, or can be found by impulse or step response experiments). Systematic methods for finding $\{P(\omega_i), \epsilon_i\}_{i=1}^{N}$ with the desired properties *in* Theorem 2.1 will be the main focus of the remainder of this paper.

## 3. SCHROEDER-PHASED INPUT DESIGN

A signal processing diagram is given in Fig. 1 for the nonparametric frequency domain identification scheme to be discussed in this section.

Consider the periodic input design composed of a harmonically related sum of sinusoids,

$$u_s(k) = \sum_{i=1}^{n_s} \sqrt{2\alpha_i} \cos(\omega_i kT + \phi_i) \qquad (3.1)$$

where $\omega_i = 2\pi i/T_p$, $T_p/T = N_s$, $n_s \leq N_s/2$. The total power is normalized as,

$$\sum_{i=1}^{n_s} \alpha_i = 1 \tag{3.2}$$

where the power in each component $\{\alpha_i > 0, i = 1, ...n_s\}$ is assumed specified. In order to minimize peaking in time domain the sinusoids are phased according to Schroeder [24] as,

$$\phi_i = 2\pi \sum_{j=1}^{i} j\alpha_j \tag{3.3}$$

(Here, a slightly modified form of the Schroeder phase is used in (3.3), as derived in Young and Patton [27]). For example, assuming equal powers $\alpha_i = 1/n_s$, gives, $\phi_i = \frac{\pi}{n_s}(i^2 + i)$. Linear terms which appear in $i$ can always be dropped since they correspond to pure time shifts. The practical use of Schroeder-phased harmonic signals for system identification has been described by Flower, Knott and Forge [11], and its superiority to sine-sweep profiles for identification of helicopter dynamics has been demonstrated in Young and Patton [27].

For technical reasons, the following assumption will be made,

Assumption 3 The system is driven by Schroeder-phased sinusoidal input (3.1), and allowed to reach steady-state before experimental data is taken.                    ▥

At steady-state, the plant response to $u_s$ is denoted as $y_s$ and is given by,

$$y_s(k) = \sum_{i=1}^{n_s} \left( b_i \sqrt{2\alpha_i} \cos(\omega_i kT + \phi_i) - a_i \sqrt{2\alpha_i} \sin(\omega_i kT + \phi_i) \right) + v(k) \tag{3.4}$$

where,

$$a_i = \%\{P^*(e^{-j\omega_i T})\}, \quad b_i = \Re\{P^*(e^{-j\omega_i T})\} \tag{3.5}$$

For notational convenience, the index $k$ starts from O in (3.4) even though we are in steady-state. Since the goal is to estimate the quantities $a_i$ and $b_i$ it is convenient to collect these quantities in a single vector $\theta$ defined as follows,

$$\theta = [a^T, b^T]^T \tag{3.6a}$$

$$a = [a_1, ..., a_{n_s}]^T; \quad b = [b_1, ..., b_{n_s}]^T \tag{3.6b}$$

In order to "whiten" the effect of the noise in (3.4), the time domain input $u_s$ and output $y_s$ will be inverse filtered by $\overline{W}$ to give filtered signals $\tilde{u}_s$ and $\tilde{y}_s$ as follows,

$$\overline{W}(z^{-1})\tilde{y}_s(k) = y_s(k); \quad \overline{W}(z^{-1})\tilde{u}_s(k) = u_s(k) \tag{3.7}$$

Since the frequencies in $u_s$ *are* harmonically related, both the input $\tilde{u}_s$ and deterministic part of the output $\tilde{y}_s$ at steady-state will be periodic with period $T_p$. Assume that m periods of filtered input/output data $\tilde{u}_s, \tilde{y}_s$ are collected at steady-state. Denote the output data from the $\ell$th period as,

$$\tilde{y}_s^\ell(k) = \tilde{y}_s(k + (\ell - 1)N_s) \tag{3.8}$$

6

for $k = 0, ..., N_s - 1$ and $\ell = 1, ..., m$.

**Remark 3.1** It is noted that when inverse filtering by $\overline{W}$ is used, the steady-state assumption in Assumption 3 requires that the filter transient settles out in addition to the plant transient. ∎

Frequency domain estimates $\hat{P}, \hat{a}_i, \hat{b}_i$ are now constructed by taking DFT's on the filtered time-domain data,

*DFT Frequency Domain Estimator*

$$\hat{P}(\omega_i) = \frac{\frac{1}{m} \sum_{\ell=1}^{m} \tilde{Y}_s^\ell(\omega_i)}{\tilde{U}_s(\omega_i)} \tag{3.9}$$

$$\hat{a}_i = \Im\{\hat{P}(\omega_i)\}, \quad \hat{b}_i = \Re\{\hat{P}(\omega_i)\} \tag{3.10}$$

where,

$$\tilde{Y}_s^\ell(\omega_i) = \frac{1}{N_s} \sum_{k=0}^{N_s-1} \tilde{y}_s^\ell(k) e^{-j\omega_i kT}; \quad \tilde{U}_s(\omega_i) = \frac{1}{N_s} \sum_{k=0}^{N_s-1} \tilde{u}_s(k) e^{-j\omega_i kT} \tag{3.11}$$

**Remark 3.2** It is emphasized that the DFT is evaluated precisely on the points of support of the Schroeder-phased input (3.1). The use of an FFT to implement (3.11) requires using the full number of sinusoids $(n_s = N_s/2)$ in the sum (3.1), choosing the frequency separation in the input design $2\pi/(N_s T)$ such that the number of samples $N_s$ in one period of $u_s$ is some power of 2. ∎

**Remark 3.3** The DFT estimator is conveniently computed recursively in the number of data windows $m$ since one can keep track of the running sum, $\sum_{\ell=1}^{m} \tilde{Y}_s^\ell(\omega_i)$ where each $\tilde{Y}_s^\ell$ is computed using an FFT of fixed size. ∎

It has been assumed that $W = \sigma\overline{W}$, where $\overline{W}$ is assumed known, and $\sigma$ may be either known or unknown. If $\sigma$ is unknown, it can be estimated as follows,

*Noise Variance Estimator*

$$\hat{\sigma}^2 = \frac{\sum_{\ell=1}^{m} \sum_{i=1}^{n_s} |\tilde{Y}_s^\ell(\omega_i) - \overline{Y}(\omega_i)|^2}{N_s(mN_s - 2n_s)} \tag{3.12}$$

where,

$$\overline{Y}(\omega_i) = \frac{1}{m} \sum_{\ell=1}^{m} \tilde{Y}_s^\ell(\omega_i) \tag{3.13}$$

## 4. STATISTICAL ANALYSIS

### 4.1 General Results

A detailed statistical analysis of frequency domain estimator (3.9) is given in Appendix A for the case $W = \sigma \cdot I$ and in Appendix B for the more general noise case $W = \sigma\overline{W}$. The results will be summarized below for convenience.

**To** aid subsequent discussion, a vector $\hat{\theta}$ of estimated quantities $\hat{a}_i$ and $\hat{b}_i$ in (3.10) is defined as follows,

$$\hat{\theta} = [\hat{a}_1, ..., \hat{a}_{n_s}, \hat{b}_1, ..., \hat{b}_{n_s}]^T \tag{4.1}$$

### Theorem 4.1

Assume that the Schroeder-phased sinusoidal input $u_s$ defined in (3.1) is applied to exponentially stable plant $P^*$ (z-l) (Assumption 1), giving rise to the steady-state output $y_s$ defined in (3.4)(3.5). Let the measurement noise coloring filter be given by $W = \sigma \overline{W}$ (Assumption 2) and implement inverse filtering of $u_s$ and $y_s$ by $\overline{W}$ (cf., *(3.7))* giving rise to filtered input $\tilde{u}_s$ and filtered output $\tilde{y}_s$. Let frequency domain estimates $\hat{P}$ defined in (3.9) and $\hat{\theta}$ defined in (3.10)(4.1), be computed based on m >1 periods of the filtered steady-state data $\{\tilde{y}_s^\ell\}_{\ell=1}^m$ in response to the Schroeder-phased input (3.1) (Assumption 3). Then,

**4.1a** If $\sigma$ is *known,* the exact error probability y distributions are given as,

$$\frac{|P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)|^2}{\sigma^2 c_{ii}} \sim \chi^2(2); \quad \theta - \hat{\theta} \sim N(0, \Sigma) \tag{4.2}$$

$$\Sigma = \sigma^2 \begin{pmatrix} C & 0 \\ 0 & C \end{pmatrix} \tag{4.3}$$

$$\mathbf{C} = diag[c_{11}, ..., c_{n_s n_s}]; \; c_{ii} = |\overline{W}(e^{-j\omega_i T})|^2 / (\alpha_i m N_s) \tag{4.4}$$

where $X^2(v)$ denotes a Chi-Squared distribution with v degreees of freedom.

**4.1b** If $\sigma$ is *unknown,* and estimated using (3,12), the exact error probability distributions are given as,

$$(mN_s - 2n_s)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(mN_s - 2n_s) \tag{4.5}$$

$$\frac{|P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)|^2}{2\hat{\sigma}^2 c_{ii}} \sim F(2, mN_s - 2n_s) \tag{4.6}$$

$$\frac{\hat{a}_i - a_i}{\hat{\sigma}\sqrt{c_{ii}}} \sim t(mN_s - 2n_s); \quad \frac{\hat{b}_i - b_i}{\hat{\sigma}\sqrt{c_{ii}}} \sim t(mN_s - 2n_s) \tag{4.7}$$

where $F(\nu_1, \nu_2)$ denotes a Fisher distribution with $\nu_1$ and $V_2$ degrees of freedom, and $t(\nu)$ denotes a Student t distribution with $\nu$ degrees of freedom.

**Proof:** The key step is to first prove that the DFT estimator (3.9) is the Gauss-Markov estimator for the stated problem. The results (4.2)-(4.7) can then be developed from existing statistical theory. Although straightforward in principle, the details are tedious and have been delegated to Theorems A.1, A.2, and B.1 in the Appendices. •

The following corollary to Theorem 4.1 is useful when statistical confidence regions are desired. It follows directly from Theorem 4.1 and is stated without proof.

## Corollary 4.1

Under the conditions of Theorem 4.1, the $(1 - a)$. 100% confidence bounds associated with the DFT estimates (3.9)(3.10) are summarized below,

$$|P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)|^2 \leq \begin{cases} c_{ii}\sigma^2 \chi^2_{1-\alpha}(2) & \text{for } \sigma^2 \text{ known} \\ c_{ii}\hat{\sigma}^2 2F_{1-\alpha}(2, mN_s - 2n_s) & \text{for } \sigma^2 \text{ estimated by } \hat{\sigma}^2 \end{cases} \quad (4.8)$$

$$|\hat{a}_i - a_i|^2 \leq \begin{cases} c_{ii}\sigma^2 \eta^2_{1-\alpha} & \text{for } \sigma^2 \text{ known} \\ c_{ii}\hat{\sigma}^2 t_{1-\alpha}(mN_s - 2n_s) & \text{for } \sigma^2 \text{ estimated by } \hat{\sigma}^2 \end{cases} \quad (4.9)$$

$$|\hat{b}_i - b_i|^2 \leq \begin{cases} c_{ii}\sigma^2 \eta^2_{1-\alpha} & \text{for } \sigma^2 \text{ known} \\ c_{ii}\hat{\sigma}^2 t^2_{1-\alpha}(mN_s - 2n_s) & \text{for } \sigma^2 \text{ estimated by } \hat{\sigma}^2 \end{cases} \quad (4,10)$$

$$c_{ii} = |\overline{W}(e^{-j\omega_i T})|^2 / (\alpha_i mN_s) \quad (4.11)$$

where $\eta_{1-\alpha}, t^2_{1-\alpha}(\nu), \chi^2_{1-\alpha}(\nu)$ and $F_{1-\alpha}(\nu_1, \nu_2)$ denote the $(1 - \alpha) \cdot 100$ percentiles for the Gaussian distribution, the Student t distribution with v degrees of freedom, the Chi-Squared distribution with $\nu$ degrees of freedom, and the Fisher distribution with $\nu_1$ over $\nu_2$ degrees of freedom, respectively. •

**Remark 4.1** To avoid confusion, it is pointed out that percentiles for symmetric densities (Gaussian and Student) used in this paper are assumed to be two-sided i.e., for $x$ Gaussian, the percentile $\eta_{1-\alpha}$ is defined as,

$$Prob\{-\eta_{1-\alpha} \leq x \leq \eta_{1-\alpha}\} = 1 - \alpha$$

•

Several important properties of the frequency domain estimates (proved in Appendices A and B) are summarized below,

**P.1** $\hat{P}(\omega_i), \hat{a}_i,$ and $\hat{b}_i,$ are unbiased and consistent estimators of $P^*(e^{-j\omega_i T}), \Im\{P^*(e^{-j\omega_i T})\},$ and $\Re\{P^*(e^{-j\omega_i T})\},$ respectively for $i = 1, \ldots . n_s.$

**P.2** $\hat{a}_i$ is statistically independent of $\hat{a}_j$ for $i \neq j$

**P.3** $\hat{b}_i$ is statistically independent of $\hat{b}_j$ for $i \neq j$

**P.4** $\hat{a}_i$ is statistically independent of $\hat{b}_j$ for all $i$ and $j$

**P.5** $\hat{P}(\omega_i)$ is statistically independent of $\hat{P}(\omega_j)$ for $i \neq j$

**P.6** $\hat{\sigma}^2$ is an unbiased and consistent estimator of $\sigma^2$

For the purpose of visualization, the estimate $\hat{P}(\omega_i)$ and its confidence region are depicted in a Nyquist plot in Fig. 2. Here the confidence region for the case of $\sigma^2$ estimated by $\hat{\sigma}^2$ is seen as a perfect circle centered at $\hat{P}(\omega_i) = \hat{b}_i + j\hat{a}_i$ of radius $\epsilon_i^2$ where from (4.8),

$$\epsilon_i^2 = \frac{\hat{\sigma}^2 |\overline{W}(e^{-j\omega_i T})|^2 2F_{1-\alpha}(2, mN_s - 2n_s)}{\alpha_i mN_s} \quad (4.12)$$

9

Noting that $F_{1-\alpha}(2,\nu)$ is bounded as $\nu$ becomes large (e.g., $F_{1-\alpha}(2,\nu) \leq 9$ for $1-\alpha = .999$ and $\nu > 30$, [7]) the uncertainty region increases with noise-to-signal ratio $\hat{\sigma}^2 \overline{W}/\alpha_i$ and decreases with the amount of measurement data $mN_s$.

**Remark 4.2** Under the conditions stated in Theorem 4.1, the error distributions (4.2)-(4.7) and confidence regions (4.8)-(4.11) are exact. However, they become approximate if the conditions are violated, i.e.,

1. The system has not reached steady-state before the data is taken

2, The noise v(k) is not Gaussian

3. Inverse filtering by $\overline{W}$ is omitted

The effect of 1. is to create a bias in the estimate, and the previous expressions must be correspondingly modified. In contrast, it is noted in Remarks A. 1 and B. 1 of the Appendix that the effect of 2. and 3. are mild and may be violated in practice while maintaining reasonable results. ∎

## *4.2* **Comparison With Other Input Designs**

Several key properties of the Schroeder-phased input design have already been summarized in the introductory remarks of Sect, 1. It is emphasized that these properties arise from the special structure of the Schroeder-phase input and are not true for more general choices of input design [8] [16] [19]. For example, the reader is referred to error estimates on page 156 of Ljung [19], where it can be seen that none of the properties listed in Sect. 1 are generically true for bounded quasi-stationary input designs on finite data sets.

## 5, STATISTICAL PLANT SET ESTIMATION

### 5.1 Basic Algorithm

An algorithm for plant set estimation based on the Schroeder-phased sinusoidal input design is given by the following sequence of steps,

1. Apply Schroeder-phased sinusoidal input design with full number of sinusoids (i.e., $n_s = N./2$) to plant. Use DFT frequency domain estimator (3.9) to find noisy frequency domain data $\{\hat{P}(\omega_i)\}_{i=1}^{n_s}$ on the uniform grid on the unit circle $\omega_i = i\Delta_g$, $i = 1,\dots, n_s$ with grid spacing $\Delta_g = \omega_{i+1} - \omega_i = \frac{\pi}{Tn_s}$.

2. Set N = $n_s$ in Theorem 2.1, specify $M$ and $\rho$, and specify stable nominal parametric plant estimate $P^o(z^{-1})$, (it is useful in practice to determine PO from a frequency weighted curve fit to the data [3]).

3. Specify $1-\kappa$ and compute $\epsilon_i$, $i = 1,\dots,n_s$ in Theorem 2.1 by the $(1-\kappa)\cdot 100\%$ confidence bounds in Corollary 4.1 (i.e., choose $\epsilon_i^2 = c_{ii}\sigma^2\chi_{1-\kappa}^2(2)$ for $\sigma^2$ known; or choose $\epsilon_i^2 = c_{ii}\hat{\sigma}^2 2F_{1-\kappa}(2, mN_s - 2n_s)$ for $\sigma^2$ estimated by $\hat{\sigma}^2$ in (3.12)). $(1-\kappa)\cdot 100\%$ DC error $\epsilon_0$.

4. Calculate statistical overbound $\ell_A^{1-\alpha}$ from Theorem 2.1, equation (2.8), where the confidence factors are $1-\alpha = (1-\kappa)^{n_s}$ for the case where $\sigma^2$ is known; or $1-\alpha = 1 - n_s\kappa$ for the case where $\sigma^2$ is estimated by $\hat{\sigma}^2$.

Due to step 3. above, the event (2.6) is satisfied with at least probability $1-\kappa$ at each grid point. In the case where $\sigma$ is known, these events are statistically independent from one grid point to the next by property P.5 of Sect. 4. Hence all conditions of Theorem 2.1 are satisfied, and $\ell_A^{1-\alpha}(\omega)$ is an overbound on the additive uncertainty with statistical confidence $1 - \alpha = (1 - \kappa)^{n_s}$.

Alternatively, in the case where $\hat{\sigma}$ is estimated, the events (2.6) are not jointly independent, (i.e., $\epsilon_i^2 = c_{ii}\hat{\sigma}^2 2F_{1-\kappa}(2, mN_s - 2n_s)$ now depends on 62), and from Theorem 2.1, the quantity $\ell_A^{1-\alpha}(\omega)$ is an overbound on the additive uncertainty with statistical Confidence $1 - \alpha = 1 - n_s\kappa$.

## 5.2 Maximum Modulus Student t-Intervals

For $n_s\kappa \ll 1$, the expression $1-n_s\kappa$ *is* close to $(1 - \kappa)^{n_s}$ since the two quantities are related by a first order Taylor expansion. In general, however, the expression can be somewhat conservative since it is based on the Bonferroni inequality in Theorem 2.1 which makes no assumption as to the underlying probability distributions.

Less conservative bounds can be found in the case where $\sigma^2$ is estimated by $\hat{\sigma}^2$ using information about the joint probabilities. For example, from (4.2), one can construct the Normal vector, $x \triangleq \Sigma^{-1/2}(\hat{\theta} - 6) \sim N(O, I)$ and from (4.5) one has $z \triangleq \nu\hat{\sigma}^2/\sigma^2 \sim \chi^2(\nu)$ where $v \triangleq mN_s - 2n_s$. It can also be shown using an argument identical to the proof of Theorem A.2 that $x$ and z are statistically independent. Hence, the random vector defined by the ratio $\tau \triangleq x/\sqrt{z/\nu} \in R^p$ ($p = 2n_s$ in the present case), is known to have a $p$-Multivariate Student t probability density (cf., Anderson [1], pp. 283), given by $\tau \sim f(\tau_1, \ldots, \tau_p; \nu)$ where,

$$f(\tau_1, ..., \tau_p; \nu) \triangleq \frac{\Gamma[\frac{1}{2}(\nu + p)]}{\Gamma[\frac{1}{2}\nu](\nu\pi)^{\frac{1}{2}p}} \cdot (1 + \tau^T\tau/\nu)^{-\frac{1}{2}(\nu+p)} \tag{5.1}$$

Precise simultaneous confidence regions can be determined directly from the joint probability distribution (5.1). For example, consider the Maximum Modulus Student t-Interval $u = u_{p,\nu}^\gamma$ defined by the solution to the following equation,

$$7 = \int_{-u}^{u} \cdots \int_{-u}^{u} f(\tau_1, \ldots \tau_p; \nu)d\tau_1...d\tau_p \tag{5.2}$$

Values of $u_{p,\nu}^\gamma$ have been tabulated in Hahn and Hendrickson [12]. Clearly, it follows from (5.2) that the set of simultaneous rectangular intervals given by,

$$|\hat{a}_i - a_i| \le \hat{\sigma}\sqrt{c_{ii}}\ u_{p,\nu}^{1-\alpha}; \quad |\hat{b}_i - b_i| \le \hat{\sigma}\sqrt{c_{ii}}\ u_{p,\nu}^{1-\alpha}; \quad i = 1, ..., 2n_s \tag{5.3}$$

will have an overall statistical confidence of *exactly* $1 - \alpha$. For each $i$, circumscribing the rectangular region (5.3) by a circle of radius $\epsilon_i = \hat{\sigma}\sqrt{2c_{ii}}\ u_{p,\nu}^{1-\alpha}$ ensures that all events $E_i$ in (2.6) are satisfied *simultaneously* with an overall probability greater than $1 - \alpha$. Using the maximum modulus t-intervals in this manner will generally provide less conservative confidence regions than using the Bonferroni inequality of Theorem 2.1.

11

Finally, a much simpler (although approximate) method is given which can be used for most practical purposes. It can be shown that $E[\tau] = O$, and $E[\tau\tau^T] = I \cdot \nu/(\nu - 2)$ for $\tau$ distributed as in (5.1) (Anderson [1]). This implies that the quantities in (4.7) are uncorrelated from one frequency to the next. As $\nu$ increases, the distribution (5.1) approaches joint Normal, for which uncorrelatedness implies statistical independence. Hence, in the limit as $\nu$ becomes large, one can use the approach outlined in Sect. 5.1 for the case where $\sigma^2$ is estimated by $\hat{\sigma}^2$, but this time assuming that the events (2.6) are statistically independent, i.e., letting $1 - \alpha = (1 - \kappa)^{n_s}$. Since $\nu = mN_s - 2n_s$ is typically large in the present application, this approach should give quite reasonable results.

### 5.3 Robust Control Applications

The additive uncertainty overbound $\ell_A^{1-\alpha}(\omega)$ can be used for either robust control analysis or synthesis, If used for synthesis, the additive uncertainty is typically represented as the product of a norm bounded perturbation and a parametric weighting transfer function $W_A$, i.e.,

$$\ell_A^{1-\alpha}(\omega) = \Delta(\omega) W_A(e^{-j\omega T}) \tag{5.4}$$

$$|\Delta(\omega)| \leq 1 \quad for \quad all \quad \omega \in [0, \pi/T] \tag{5.5}$$

Given $\ell_A^{1-\alpha}$, the weighting function $W_A$ in (5.4) can be found using graphical methods, or a more systematic optimal minimax method recently proposed in [23].

Given $P^o$ and $WA$, controllers can be designed using p-synthesis and $H_\infty$ software **[2][9]** to provide stability and a prescribed degree of performance *for all plants in the uncertainty set (2.3)*. This overall procedure involves a Tustin transformation of $P^o$ and $W_A$ into the w-Domain, a robust control design using the software in [2] or [9], and then an inverse Tustin transformation of the final compensator for implementation in the z-domain. Since by definition the true plant lies in the uncertainty set (2.3) with probability $1 - \alpha$, this approach assures that with $1 - \alpha\%$ confidence the robust controller will work as designed when applied to the true system. Since the size of the uncertainty set (2.3) will decrease with a reduction in the confidence factor $1 - \alpha$, this approach provides a clear trade-off between performance and risk.

As an alternative to using Theorem 2.1, one can use the Schroeder-phased multisinusoidal input design to support the recent method of Helmicki, Jacobson and Nett [15]. Their scheme requires frequency domain data of the form $\{\hat{P}(\omega_i), \epsilon_i\}$ where $\epsilon_i$ denotes a bound on the error in $\hat{P}(\omega_i)$. This is precisely the data available from the multisinusoidal input design. Replacing hard bounds (found their paper using one-at-a-time sinusoidal excitation) with soft bounds derived here, allows all of their results to be interpreted in a rigorous statistical framework. This also has the advantage that the Schroeder-phased sinusoids are applied in parallel, and that the statistical bounds go to zero asymptotically at the frequency grid points, while the hard bounds derived in [15] do not.

## 6. CONCLUSIONS

A frequency domain identification methodology was introduced using a Schroeder-phased multisinusoidal input design. Exact expressions were derived for the error distributions under the assumption that data is gathered while the system is in steady-state. The error distributions were shown to be plant independent, unbiased, consistent, and statistically independent from one frequency to the next.

Using the special properties of the Schroeder based frequency domain estimates and certain a-priori information as to the system damping, a method was developed to characterize the plant set to a precise statistical confidence. In particular, the plant set is characterized by a statistical bound $\ell_A^{1-\alpha}$ on the additive uncertainty, which can be used to replace hard bounds presently used exclusively in the robust control literature.

If a robust controller is designed to provide some specified level of stability or performance for all plants lying in the additive uncertainty set defined by $\ell_A^{1-\alpha}$, then, with probability $1 - \alpha$ the controller will work as designed *when applied to the true system.* This effects a '(marriage" between identification and control disciplines such that noise is allowed to '(average out" over time, giving potentially less conservative uncertainty descriptions compared to purely deterministic treatments.

## REFERENCES

[1] T.W. Anderson, *An Introduction to Multivariate Statistical Analysis.* 2nd ed., Wiley, New York, 1984.

[2] G.J. Balas, J.C. Doyle, K. Glover, A.K. Packard, R. Smith, *H-Infinity and Mu Control Analysis : Mu-Tools Manual* (beta test version), September 1990.

[3] D.S. Bayard, F.Y. Hadeagh, Y. Yam, R.E. Scheid, E. Mettler, M.H. Milman, "Automated On-Orbit Frequency Domain Identification for Large Space Structures," Automatic, Vol. 27, No. 6, pp. 931-946, November 1991.

[4] D.S. Bayard, "Multivariable Frequency Domain Identification via 2-Norm Minimization," Proc. American Control Conference, Chicago, Illinois, June 1992.

[5] D.S. Bayard, "Statistical Additive Uncertainty Bounds Using Schroeder-Phased Input Design," Jet Propulsion Laboratory, Internal Document, JPL D-8146, January 1991; see also, Proc. American Control Conference, Chicago, Illinois, June 1992.

[6] D.S. Bayard, Y. Yam, E. Mettler, "A criterion for joint optimization of identification and robust control," IEEE Transactions on Automatic Control, *Special Mini-Issue on System Identification for Control Design,* Vol. 37, No. 7, pp. 986-991, July 1992.

[7] R.M. Bethea, B.S. Duran and T.L. Boullion, *Statistical Methods for Engineers and Scientists.* Marcel Dekker, New York, 1975.

[8] D.R. Brillinger, *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco, 1981.

[9] R.Y. Chiangand M.G. Safonov, *Robust-Control Toolbox: Users* Guide The MathWorks, Inc., Sout Natick, MA, June 1988.

[10] W. Feller, *An Introduction to Probability y Theory and its Applications.* 3rd ed., Wiley, New York, 1968.

[11] J.O. Flower, G.F. Knott, and S.C. Forge, "Application of Schroeder-phased harmonic signals to practical identification," J. Dynamic Systems, pp. 69-73, Vol. 16, Feb. 1978.

[12] Hahn and Hendrickson, "A table of percentage points of the distribution of the largest absolute value of k Student t variates and its applications," Biometrika, **vol.** 58, pp. 323-332, 1971.

[13] G.C. Goodwin and R.L. Payne, *Dynamic System Identification: Experiment Design and Data Analysis,* Academic Press, New York.

[14] G.C. Goodwin and M.E. Salgado, "Quantification of uncertainty in estimation using an embedding principle," Proc. American Control Conference, Pittsburg, PA, June 1989.

[15] A. Helmicki, C.A. Jacobson, and C.M. Nett, "Identification in $H_\infty$: A robustly convergent nonlinear algorithm," Proc. ACC, pages 1428-1434, Pittsburgh, PA, June 1989.

[16] G.M. Jenkins and D.G. Watts, *Spectral Analysis and its Applications.* Holden-Day, San Francisco, 1968.

[17] R.L. Kosut, "On-line identification and control tuning of large space structures," Proc. Fifth Yale Conference on Adaptive Systems Theory, Yale University, May 1987.

[18] R.L. Kosut, M. Lau, and S. Boyd, "Identification of systems with parametric and non-pararnetric uncertainty," Proc. American Control Conf., pp. 2412-2417, San Diego, 1990.

[19] L. Ljung, *System Identification: Theory for the User.* Prentice-Hall, Englewood Cliffs, New Jersey, 1987.

[20] A.V. Oppenheim and R.W. Schafer, *Digital* Signal Processing. Prentice-Hall, Englewood Cliffs, N. J., 1975.

[21] A. Papoulis, *Probability, Random Variables, and Stochastic Processes,* McGraw-Hill, New York, 1984.

[22] D.E. Rivera, J.F. Pollard, L.E. Sterman, and C.E. Garcia, "An industrial perspective on control-relevent identification," Proc. American Control Conference, pp. 2406-2411, San Diego, CA, 1990.

[23] R.E. Scheid, D.S. Bayard, Y. Yam, "A Linear Programming Approach to Characterizing Norm Bounded Uncertainty From Experimental Data," American Control Conference, Boston, MA, pp. 1956-1958, June 1991.

[24] M.R. Schroeder, "Synthesis of low peak-factor signals and binary sequences of low auto-correlation," IEEE Trans. Information Theory, pp. 85-89, Vol. 16, Jan. 1970.

[25] R.S. Smith and J.C. Doyle, "Model validation: A connection between robust control and identification," Proc. American Control Conference, Pittsburg, PA, June 21-23, 1989.

[26] R. C. Younceand C. E. Rohrs, " Identification with non-parametric uncertainty, " Proc. Int. Conf. on Circuits and Systems, 1990,

[27] P. Young and R.J. Patton, "Comparison of test signals for aircraft frequency domain identification," AIAA J. Guidance, Dynamics and Control, vol. 13, no. 3, pp. 430-438, May-June 1990.

## APPENDIX A

The analysis in Appendix A is self-contained, and inapplicable to systems for which the output noise coloring filter in Fig. 1 is assumed to be a constant $\sigma$ or equivalently, $\overline{W} = 1$.

Consider the periodic input design composed of a harmonically related sum of sinusoids,

$$u_s(k) = \sum_{i=1}^{n_s} \sqrt{2\alpha_i} \cos(\omega_i kT + \phi_i) \tag{A1}$$

where $\omega_i = 2\pi i/T_p, T_p/T = N_s, n_s \le N_s/2$. While the phase terms $\phi_i$ should be chosen in practice to minimize peaking in time (i.e., via Schroeder), the particular choice of phase in (A. 1 ) will not effect the results in this section.

Assuming that the plant is exponentially stable, the system will reach steady-state, at which time the plant response to $u_s$ is denoted as $y_s$ and is given by,

$$y_s(k) = \overline{y}(k) + v(k) \tag{A.2}$$

where,

$$\overline{y}(k) = \sum_{i=1}^{n_s} b_i \sqrt{2\alpha_i} \cos(\omega_i kT + \phi_i) - a_i \sqrt{2\alpha_i} \sin(\omega_i kT + \phi_i) \tag{A.3}$$

$$a_i = \Im\{P^*(e^{-j\omega_i T})\}, \quad b_i = \Re\{P^*(e^{-j\omega_i T})\} \tag{A.4}$$

Since the goal is to estimate the quantities $a_i$ and $b_i$ it is convenient to collect these quantities in a single vector $\theta$ defined as follows,

$$\theta = [a^T, b^T]^T \tag{A.5a}$$

$$a = [a_1, ..., a_{n_s}]^T, \quad b = [b_1, ..., b_{n_s}]^T \tag{A.5b}$$

For notational convenience, the index $k$ starts from O in (A.2) even though we are in steady-state. Since the frequencies in $u_s$ are harmonically related, both the input $u_s$ and deterministic part of the output at steady-state $\overline{y}$ will be periodic with period $T_p$.

Assume that $m$ periods of data are collected at steady-state. Denote the output data from the $\ell$th period as,

$$y_s^\ell(k) = y_s(k + (\ell - 1)N_s) \tag{A.6}$$

for $k = 0, ..., N_s - 1$ and $\ell = 1, ..., m$. Frequency domain estimates $\hat{P}(\omega_i), \hat{a}_i, \hat{b}_i$ are now constructed from the data.

$$\hat{P}(\omega_i) = \frac{\frac{1}{m}\sum_{\ell=1}^{m} Y_s^\ell(\omega_i)}{U_s(\omega_i)} \tag{A.7}$$

$$\hat{a}_i = \Im\{\hat{P}(\omega_i)\}, \quad \hat{b}_i = \Re\{\hat{P}(\omega_i)\} \tag{A.8}$$

where,

$$Y_s^\ell(\omega_i) = \frac{1}{N_s} \sum_{k=0}^{N_s-1} y_s^\ell(k)e^{-j\omega_i kT}; \quad U_s(\omega_i) = \frac{1}{N_s} \sum_{k=0}^{N_s-1} u_s(k)e^{-j\omega_i kT} \tag{A.9}$$

To aid subsequent discussion, a vector $\hat{\theta}$ of estimated quantities $\hat{a}_i$ and $\hat{b}_i$ in (A.8) is defined as follows,

$$\hat{\theta} = [\hat{a}^T, \hat{b}^T]^T \tag{A.10a}$$

$$\hat{a} = [\hat{a}_1, ..., \hat{a}_{n_s}]^T, \hat{b} = [\hat{b}_1, ..., \hat{b}_{n_s}]^T \tag{A.10b}$$

The analysis to follow will treat two cases, depending on whether or not $\sigma$ is known a-priori.

## CASE 1: $\sigma$ KNOWN

Consider the following lemma.

### Lemma A.1

If $A$ is an idempotent matrix (i.e., $A^2 = A$), and $x \sim N(0, I)$, then $X^T A x$ has a $\chi^2(r)$ distribution, where $r = trace(A) = rank(A)$.

**Proof: see** Goodwin and Payne [13], pg. 217. ∎

An important result follows concerning the statistical properties of the estimates, $\hat{P}, \hat{a}_i$, and $\hat{b}_i$.

### Theorem A.1

Let frequency domain estimates $\hat{P}$ defined in (A.7) and $\hat{\theta}$ defined in (A.8)(A.10), be computed based on m periods of the steady-state data $\{y_s^\ell\}_{\ell=1}^m$ (A.2)(A.6) in response to the Schroeder-phased input (Al). Let the measurement noise $v(k)$ in (A.2) be white Gaussian i.e., colored by filter $W = \sigma$ where CT is a known constant. Furthermore, let $\theta$ be the vector of real and imaginary parts of $P^*$ as defined in (A.5) and let $\hat{\theta}$ be the vector of real and imaginary parts of $\dot{P}$ as defined in (A.10). Then the statistics of $\dot{P}$ and $\hat{\theta}$ are given as,

$$\frac{|P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)|^2}{\sigma^2 c_{ii}} \sim \chi^2(2) \tag{A11}$$

$$\hat{\theta} \sim \mathbf{Iv(e, \Sigma)} \tag{A.12}$$

where,

$$\Sigma = \sigma^2 \begin{pmatrix} C & 0 \\ 0 & C \end{pmatrix} \tag{A.13}$$

$$\mathbf{C} = diag[c_{11}, ..., c_{n_s n_s}], c_{ii} = 1/(\alpha_i m N_s) \tag{A.14}$$

and where $X^2(v)$ denotes a Chi-Squared distribution with v degrees of freedom.

16

**Proof:** Assume for the moment that we treat (A.2)(A.3) as a regression equation and apply a Gauss-Markov formulation to estimating the parameters $a_i$ and $b_i$. This gives the matrix system,

$$y = \overline{y} + v = H\theta + v \tag{A.15}$$

where $y = [y_s(0), \ldots, y_s(mN_s - 1)]^T$, $\overline{y} = [\overline{y}(0), \ldots, \overline{y}(mN_s-1)]^T$, $v = [v(0), \ldots, v(mN_s-1)]^T$, $E[v] = 0$, $Cov[v] = \sigma^2 \cdot I$, and $H$ can be written in partitioned form as,

$$H = [H_a | H_b] \tag{A.16}$$

$$H_a = -\sqrt{2} \begin{pmatrix} \sqrt{\alpha_1}\sin(\omega_1 0T + \phi_1) & \cdots & \sqrt{\alpha_{n_s}}\sin(\omega_{n_s} 0T + \phi_{n_s}) \\ \vdots & \ddots & \vdots \\ \sqrt{\alpha_1}\sin(\omega_1(mN_s - 1)T + \phi_1) & \cdots & \sqrt{\alpha_{n_s}}\sin(\omega_{n_s}(mN_s - 1)T + \phi_{n_s}) \end{pmatrix}$$

$$H_b = \sqrt{2} \begin{pmatrix} \sqrt{\alpha_1}\cos(\omega_1 0T + \phi_1) & \cdots & \sqrt{\alpha_{n_s}}\cos(\omega_{n_s} 0T + \phi_{n_s}) \\ \vdots & \ddots & \vdots \\ \sqrt{\alpha_1}\cos(\omega_1(mN_s - 1)T + \phi_1) & \cdots & \sqrt{\alpha_{n_s}}\cos(\omega_{n_s}(mN_s - 1)T + \phi_{n_s}) \end{pmatrix}$$

Using Normal theory, the Gauss-Markov estimate $\theta^*$ is given by [13],

$$\theta^* = (H^T H)^{-1} H^T y \tag{A.17}$$

Furthermore, it is well known that $\theta^*$ has statistics [13],

$$\theta^* \sim N(\theta, \Sigma^*) \tag{A.18}$$

$$\Sigma^* = \sigma^2 (H^T H)^{-1} \tag{A.19}$$

Using the partitioned matrix (A.16) and the fact that $H_a^T H_b = H_b^T H_a = O$ gives after some algebra,

$$H^T H = \begin{pmatrix} H_a^T H_a & 0 \\ 0 & H_b^T H_b \end{pmatrix} = \begin{pmatrix} C^{-1} & 0 \\ 0 & C^{-1} \end{pmatrix} \tag{A.20}$$

where,

$$C = diag[c_{11}, \ldots, c_{n_s\, n_s}], \quad c_{ii} = 1/(\alpha_i mN_s) \tag{A.21}$$

Substituting (A.20) into the Gauss-Markov estimator (A.17) and using the partitioned not at ion

$$O^* = [(a^*)^T, (b^*)^T]^T \tag{A.22a}$$

$$a^* = [a_1^*, \ldots, a_n^*]^T, \quad b^* = [b_1^*, \ldots, b_n^*]^T \tag{A.22b}$$

gives,

$$a^* = C H_a^T y \tag{A.23a}$$

$$b^* = C H_b^T y \tag{A.23b}$$

Expanding, this can be shown to be componentwise equivalent to $a_i^* = \hat{a}_i = \Im\{\hat{P}(\omega_i)\}$ and $b_i^* = \hat{b}_i = \Re\{\hat{P}(\omega_i)\}$ where $\hat{P}$ is the frequency domain estimator defined by (A.7) (this is most efficiently done by abandoning matrix notation for this step, and demonstrating the result componentwise). Hence, $\hat{\theta} = \theta^*$ (i.e., the DFT estimator (A.7)(A.8) is the Gauss Markov estimator for this problem) and results (A.12)-(A.14) follow from (A.18)-(A.21).

To show (Al 1), construct the normalized statistically independent Gaussian variables, $r_a = (a_i - \hat{a}_i)/(\sigma\sqrt{c_{ii}}) \sim N(0, 1)$, and $r_b = (b_i - \hat{b}_i)/(\sigma\sqrt{c_{ii}}) \sim N(O, 1)$. Then,

$$\frac{|P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)|^2}{\sigma^2 c_{ii}} = \frac{(a_i - \hat{a}_i)^2 + (b_i - \hat{b}_i)^2}{\sigma^2 c_{ii}} = r_a^2 + r_b^2 \sim \chi^2(2) \qquad (A.24)$$

where the $\chi^2(2)$ distribution for $r_a^2 + r_b^2$ is deduced by invoking Lemma A. 1 with $x = [r_a, r_b]^T$ and $A = I$. $\blacksquare$

### Corollary A.1

Estimates $\hat{P}(\omega_i), \hat{a}_i$, and $\hat{b}_i$ from the DFT estimator (A.7)(A.8) are unbiased and consistent estimators of $P^*(e^{-j\omega_i T})$, $a_i$, and $b_i$, respectively, for all $i = 1, \ldots, n_s$.

**Proof:** From Theorem A.1 equation (A.12) it follows that $E[\hat{a}_i] = a_i$, $E[\hat{b}_i] = b_i$ and hence $E[\hat{P}(\omega_i)] = P^*(e^{-j\omega_i})$. Furthermore, the estimates converge in probability, since, for any $\epsilon$ and $\delta$,

$$Prob(|P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)| > \epsilon) < \delta$$

for all $m > \chi^2_{1-\delta}(2)\sigma^2/(\epsilon^2 \alpha_i N_s)$; and,

$$Prob(|\hat{a}_i - a_i| > \epsilon) < \delta, \quad Prob(|\hat{b}_i - b_i| > \epsilon) < \delta$$

for all $m > \eta^2_{1-\delta}\sigma^2/(\epsilon^2 \alpha_i N_s)$. $\blacksquare$

**Remark A.1** It is noted that even if v is not Gaussian, each of the estimates $\hat{a}_i$ and $\hat{b}_i$ and hence $\hat{P}(\omega_i)$ are linear combinations of $mN_s$ independent random variables. Hence, by the Central Limit Theorem [21], the distributions are well approximated by Gaussian distributions for $mN_s$ reasonably large. $\blacksquare$

### CASE 2: $\sigma$ UNKNOWN

In certain applications, the noise variance $\sigma^2$ may not be known a-priori. In this case, it can be estimated directly from the data by the quantity $\hat{\sigma}^2$ where,

$$\hat{\sigma}^2 = \frac{S(\hat{\theta})}{mN_s - 2n_s} \qquad (A.25)$$

$$S(\hat{\theta}) = (y - H\hat{\theta})^T(y - H\hat{\theta}) \qquad (A.26)$$

where y and *H are as* defined in the proof of Theorem A. 1.

18

**Remark A.2** Alternatively, $\hat{\sigma}^2$ in (A.25) can be conveniently computed in terms of the spectral estimates $Y_s^\ell(\omega_i)$ as follows,

$$\hat{\sigma}^2 = \frac{\sum_{\ell=1}^m \sum_{i=1}^{n_s} |Y_s^\ell(\omega_i) - \overline{Y}(\omega_i)|^2}{N_s(mN_s - 2n_s)} \qquad (A.27)$$

where,

$$\overline{Y}(\omega_i) = \frac{1}{m} \sum_{\ell=1}^m Y_s^\ell(\omega_i) \qquad (A.28)$$

The computation of (A.25) by formula (A.27) follows by noting,

$$\mathsf{s(e)} = (y - H\hat{\theta})^T(y - H\hat{\theta}) \qquad (A.29)$$

$$= \sum_{\ell=1}^m \sum_{k=0}^{N_s-1} \left( y_s^\ell(k) - \sum_{i=1}^{n_s} \left( \hat{b}_i \sqrt{2\alpha_i} \cos(\omega_i kT + \phi_i) \; \hat{a}_i \sqrt{2\alpha_i} \sin(\omega_i kT + \phi_i) \right) \right)^2 \qquad (A.30)$$

$$= \sum_{\ell=1}^m \sum_{i=1}^{n_s} |Y_s^\ell(\omega_i) - \hat{P}(\omega_i)U_s(\omega_i)|^2/N_s \qquad (A.31)$$

$$= \sum_{\ell=1}^m \sum_{i=1}^{n_s} |Y_s^\ell(\omega_i) - \overline{Y}(\omega_i)|^2/N_s \qquad (A.32)$$

where (A.30) follows by expanding (A.29) in terms of summation notation; (A.31) follows from (A.30) by using Parseval's formula [20]; and (A.32)) follows by using the definition of $\hat{P}$ from (A.7). ∎

The following result characterizes the error probabilities in the case where $\sigma$ is estimated,

## Theorem A.2

Let frequency domain estimates $\hat{P}$ defined in (A.7) and $\hat{\theta}$ defined in (A.8)(A.1O), be computed based on $m$ periods of the steady-state data $\{y_s^\ell\}_{\ell=1}^m$ (A.2)A.6) in response to the Schroeder-phased input (Al). Furthermore, let the output noise $v(k)$ in (A.2) be a stationary Gausian process with coloring filter $W = \sigma$ where u' is an unknown scalar, and estimated by $\hat{\sigma}^2$ in (A.25). Then,

**A.2.a** $\hat{\sigma}^2$ is an unbiased, consistent estimate of $\sigma^2$ and has the probability distribution,

$$(mN_s - 2n_s)\frac{\hat{\sigma}^2}{\sigma^2} \sim \chi^2(mN_s - 2n_s) \qquad (A.33)$$

**A.2.b**

$$\frac{\hat{a}_i - a_i}{\hat{\sigma}\sqrt{c_{ii}}} \sim t(mN_s - 2n_s); \quad \frac{\hat{b}_i - b_i}{\hat{\sigma}\sqrt{c_{ii}}} \sim t(mN_s - 2n_s) \qquad (A.34)$$

where $t(\nu)$ denotes a Student t distribution with v degrees of freedom;

**A.2.c**

$$\frac{|P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)|^2}{2\hat{\sigma}^2 c_{ii}} \sim F(2, mN_s - 2n_s) \tag{A.35}$$

where $F(\nu_1, \nu_2)$ denotes a Fisher distribution with $\nu_1$ and $\nu_2$ degrees of freedom.

**Proof:** This proof makes use of all notation and results developed in statement and proof of Theorem A. 1.

Define residual vectors $r_1 = y - H\hat{\theta}$ and $r_2 = \hat{\theta} - \theta$ where, as noted in the proof of Theorem A.1, $\hat{\theta}$ is the Gauss- Markov estimator for $\theta$. Vectors $r_1$ and $r_2$ are joint multivariate Normal since it can be shown that r, $=(1 - HL)v$ and $r_2 = Lv$ where $L = (H^T H)^{-1} H^T$. Furthermore, it can be verified that $Cov[r_1, r_2] = \sigma^2(I - HL)L^T = 0$, implying that $r_1$ and $r_2$ are statistically independent. It is also noted that $r_1$ and $Hr_2$ are orthogonal as Euclidean vectors since $(r_1, Hr_2) = v^T(I - HL)HLv = O$. These facts will be required for the proof.

Consider the random variable $z = S(\hat{\theta})/\sigma^2$ where $S(\hat{\theta})$ is defined in (A.26). By algebra,

$$z = \frac{S(\hat{\theta})}{\sigma^2} = \frac{(y - H\hat{\theta})^T(y - H\hat{\theta})}{\sigma^2} \tag{A.36}$$

$$= \frac{r_1^T r_1}{\sigma^2} = \frac{v^T(I - HL)v}{\sigma^2} \tag{A.37}$$

Since $I - HL$ is idempotent with rank $mN_s - 2n_s$, and $v/\sigma \sim N(O, I)$ it follows from (A.37) and Lemma A.1 that $z \sim \chi^2(mN_s - 2n_s)$. Result (A.33) follows by noting from (A.25) and (A.36) that $z = (mN_s - 2n_s)\hat{\sigma}^2/\sigma^2$. Unbiasedness and consistency of $\hat{\sigma}^2$ follow immediately from (A.33).

Consider the normalized random variable $\tilde{a} = (\hat{a}_i - a_i)/(\sigma\sqrt{c_{ii}}) \sim N(O, 1)$. Since $r_1$ is independent of $r_2$, it follows that $\tilde{a}$ is independent of z. Hence, we can construct a Student t distributed random variable by taking the ratio,

$$\frac{\tilde{a}}{\sqrt{z/(mN_s - 2n_s)}} = \frac{(\hat{a}_i - a_i)/(\sigma\sqrt{c_{ii}})}{\sqrt{S(\hat{\theta})/(\sigma\sqrt{mN_s - 2n_s})}} \sim t(mN_s - 2n_s) \tag{A.38}$$

Here, we have used the fact that $z/\sqrt{y/k} \sim t(k)$ where $x$ and $y$ are statistically independent random variables such that $x \sim N(O, 1)$ and $y \sim \chi^2(k)$ [7]. Writing (A.38) in terms of $\hat{\sigma}$ from (A.25) gives the desired result (A.34) for $\hat{a}_i$. An identical argument can be used to prove result (A.34) for $\hat{b}_i$.

To prove (A.35), consider normalized random variables $r_a = (\hat{a}_i - a_i)/(\sigma\sqrt{c_{ii}}) \sim N(O, 1)$, $r_b = (\hat{b}_i - b_i)/(\sigma\sqrt{c_{ii}}) \sim N(O, 1)$, and $z = S(\hat{\theta})/\sigma^2$. It is noted that $r_a^2 \sim X2(1)$, $r_b^2 \sim X2(1)$, $z \sim \chi^2(mN_s - 2n_s)$, and all three random variables $r_a, r_b, z$ are statistically independent by the statistical independence of $r_1$ and $r_2$. *Now,* by algebra it can be shown that

$$\frac{|P^*(e^{-j\omega_i T}) - \hat{P}(\omega_i)|^2}{2\hat{\sigma}^2 c_{ii}} = \frac{(r_a^2 + r_b^2)/2}{z/(mN_s - 2n_s)} \sim F(2, mN_s - 2n_s) \tag{A.39}$$

Here, we have used the fact that $\frac{(x/\nu_x)}{y/\nu_y} \sim F_{(\nu_x, \nu_y)}$ where $x$ and $y$ are statistically independent random variables such that $x \sim \chi^2(\nu_x)$ and $y \sim \chi^2(\nu_y)$. Equation (A.39) is the desired result (A.35). ∎

## APPENDIX B

The analysis in Appendix B draws upon the results from Appendix A, and is applicable to systems where the noise coloring filter in Fig. 1 is given by the general form $W = \sigma\overline{W}$ where $\overline{W}$ is assumed known, while $\sigma$ may be known or unknown.

In order to "whiten" the effect of the noise, the time domain input $u_s$ and output $y_s$ will be inverse filtered by $\overline{W}$ to give filtered signals $\tilde{u}_s$ and $\tilde{y}_s$ as follows,

$$\overline{W}(z^{-1})\tilde{y}_s(k) = y_s(k); \quad \overline{W}(z^{-1})\tilde{u}_s(k) = u_s(k) \tag{B.1}$$

Since the frequencies in $u_s$ are harmonically related, both the input $\tilde{u}_s$ and deterministic part of the output $\tilde{y}_s$ at steady-state will be periodic with period $T_p$. Assume that m periods of filtered input/output data $\tilde{u}_s, \tilde{y}_s$ are collected at steady-state. Denote the filtered output data from the $\ell$th period as,

$$\tilde{y}_s^\ell(k) = \tilde{y}_s(k + (\ell - 1)N_s) \tag{B.2}$$

for $k = 0,...,$ IV, $-1$ and $\ell = 1,...,m$.

### Theorem- B.1

For the general case of noise coloring filter $W(z^{-1}) = \sigma\overline{W}(z^{-1})$, consider the computation of $\hat{P}$ in (A.7) and $\hat{\sigma}^2$ in (A.27) with $U^*$ and $y_s^\ell(k)$ replaced by their steady-state filtered counterparts $\tilde{u}_s$ and $\tilde{y}_s^\ell$ in (B.1) and (B.2). Then the results of Theorems A.1 and A.2 remain true with $c_{ii}$ modified to

$$c_{ii} = |\overline{W}(e^{-j\omega_i T})|^2/\alpha_i mN_s \tag{B.3}$$

**Proof:** The noise is colored such that $v(k) = \sigma\overline{W}d(k)$. Hence by the filtering action in (B.1) and the steady-state assumption, it follows that,

$$\tilde{u}_s(k) = \sum_{i=1}^{n_s} \sqrt{2\tilde{\alpha}_i}\cos(\omega_i kT + \tilde{\phi}_i) \tag{B.4}$$

$$\tilde{y}_s(k) = \sum_{i=1}^{n_s} b_i\sqrt{2\tilde{\alpha}_i}\cos(\omega_i kT + \tilde{\phi}_i) - a_i\sqrt{2\tilde{\alpha}_i}\sin(\omega_i kT + \tilde{\phi}_i) + \tilde{v}(k) \tag{B.5}$$

where $\sqrt{2\tilde{\alpha}_i} = \sqrt{2\alpha_i}/|\overline{W}(e^{-j\omega_i T})|$, and the additive noise $\tilde{v}(k)$ is "whitened" to give $\tilde{v}(k) = \sigma d(k)$. This case is now identical to that treated Theorems A.1 and A.2 with $u_s$ replaced with $\tilde{u}_s$ and $y_s$ replaced by $\tilde{y}_s$ or equivalently, with $\alpha_i$ replaced by $\tilde{\alpha}_i = \alpha_i/|\overline{W}(e^{-j\omega_i T})|^2$. ∎

**Remark B.1** It is useful to note that if the length of each data window $T_p = TN_s$ is large compared to the time constants of $\overline{W}^{-1}$, then the inverse *filtering in Fig. 1 can be omitted with little error,* since $U_s \equiv DFT\{\tilde{u}_s\} = U_s/\overline{W}$ and $\hat{Y}_s^\ell \equiv DFT\{\tilde{y}_s^\ell\} \simeq Y_s^\ell/\overline{W}$ and thus,

$$\hat{P}(\omega_i) = \frac{\frac{1}{m}\sum_{\ell=1}^m \tilde{Y}_s^\ell(\omega_i)}{\tilde{U}_s(\omega_i)} \simeq \frac{\frac{1}{m}\sum_{\ell=1}^m Y_s^\ell(\omega_i)}{U_s(\omega_i)}$$

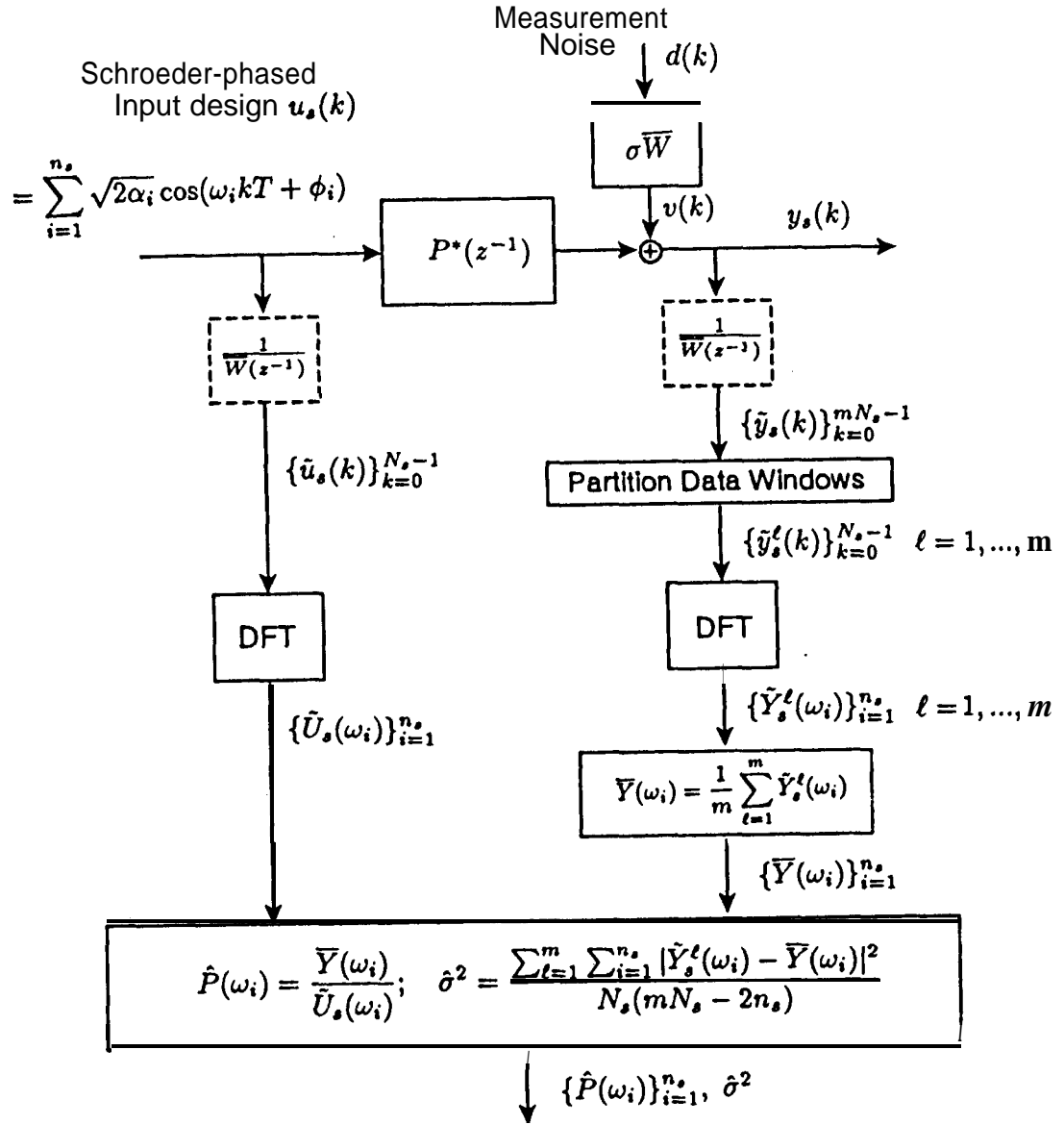This has important practical implications since the complications associated with inverse filtering are eliminated. $\bullet$

Measurement
Noise $\quad\downarrow\; d(k)$

Schroeder-phased
Input design $u_s(k)$

$$= \sum_{i=1}^{n_s} \sqrt{2\alpha_i}\cos(\omega_i kT + \phi_i)$$

$\boxed{\sigma\overline{W}}$

$P^*(z^{-1})$

$\downarrow v(k) \qquad y_s(k)$

$\boxed{\dfrac{1}{\overline{W}(z^{-1})}}$ (dashed)

$\boxed{\dfrac{1}{\overline{W}(z^{-1})}}$ (dashed)

$\{\tilde{y}_s(k)\}_{k=0}^{mN_s-1}$

$\{\tilde{u}_s(k)\}_{k=0}^{N_s-1}$

Partition Data Windows

$\{\tilde{y}_s^\ell(k)\}_{k=0}^{N_s-1} \quad \ell = 1,...,m$

DFT

DFT

$\{\tilde{U}_s(\omega_i)\}_{i=1}^{n_s}$

$\{\tilde{Y}_s^\ell(\omega_i)\}_{i=1}^{n_s} \quad \ell = 1,...,m$

$$\overline{Y}(\omega_i) = \frac{1}{m}\sum_{\ell=1}^{m}\tilde{Y}_s^\ell(\omega_i)$$

$\{\overline{Y}(\omega_i)\}_{i=1}^{n_s}$

$$\hat{P}(\omega_i) = \frac{\overline{Y}(\omega_i)}{\tilde{U}_s(\omega_i)}; \quad \hat{\sigma}^2 = \frac{\sum_{\ell=1}^{m}\sum_{i=1}^{n_s}|\tilde{Y}_s^\ell(\omega_i) - \overline{Y}(\omega_i)|^2}{N_s(mN_s - 2n_s)}$$

$\downarrow \quad \{\hat{P}(\omega_i)\}_{i=1}^{n_s}, \; \hat{\sigma}^2$

Figure 1. **Frequency domain** estimation **using Schroeder-phased sinusoid input design.**

$$\epsilon_i^2 = \frac{\hat{\sigma}^2 |\overline{W}(e^{-j\omega_i T})|^2 2F_{1-\alpha}(2, mN_s - 2n_s)}{\alpha_i mN_s}$$
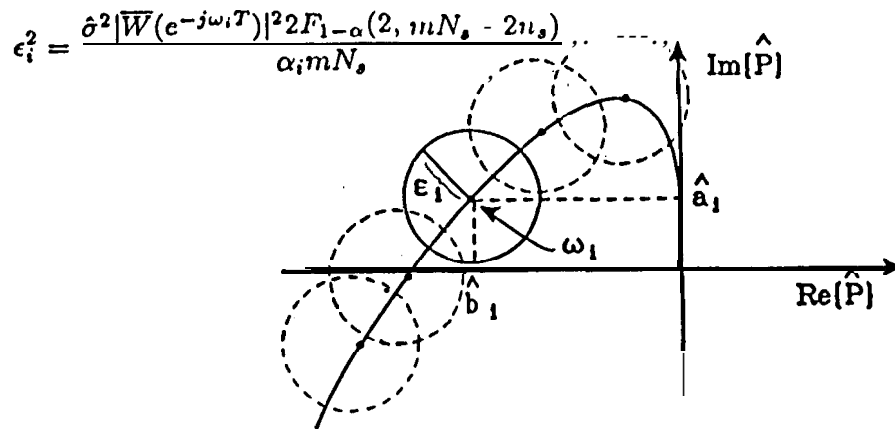


Figure 2.- Nyquist plot showing frequency domain estimate $\hat{P}(\omega_i)$ with $(i-\alpha) \cdot 100\%$ confidence radius $\epsilon_i$.
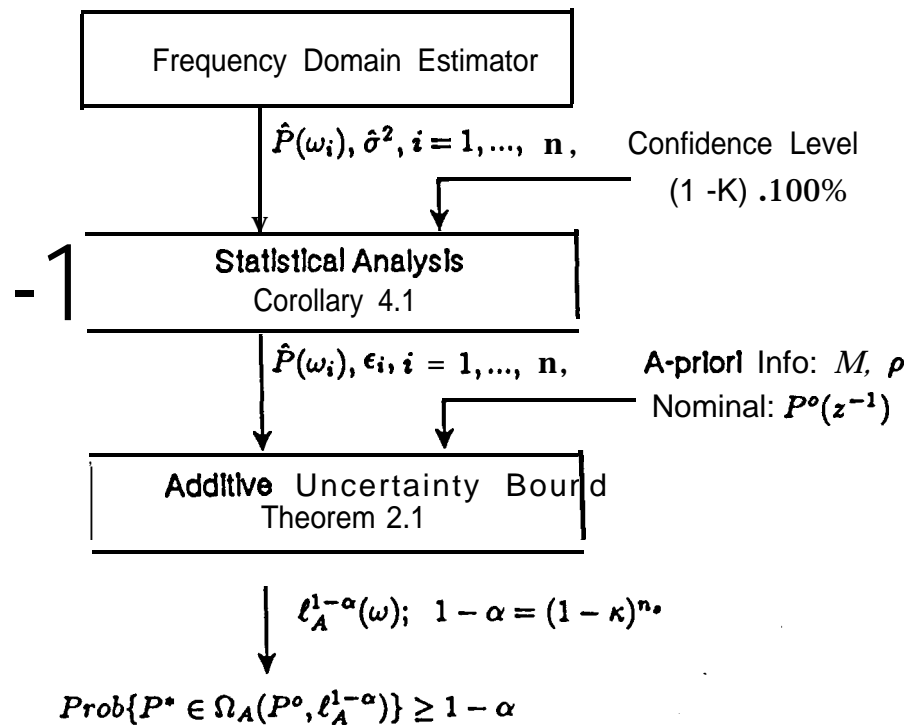


Figure 3. Sequence for computing additive uncertainty overbound $\ell_A^{1-\alpha}(\omega)$ having statistical confidence (1 -a)+ '1OO%.